Discrete Mathematics 311 (2011) 349–361

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Proppian random walks in $\mathbb{Z}$

Juliana Freire [a,*], Joel Spencer [b,1]

[a] PUC-Rio, Depto. de Matemática, Rua Marques de Sao Vicente 225, Ed. Cardeal Leme, 8o. Andar, Gávea, 22451-900, Rio de Janeiro, RJ, Brazil

[b] New York University, 251 Mercer Street, New York, NY 10012, United States

ARTICLE INFO

Article history:
Received 18 August 2009
Received in revised form 26 October 2010
Accepted 1 November 2010
Available online 27 November 2010

Keywords:
Propp machine
Random walk
Derandomization

ABSTRACT

The Propp Machine is a deterministic process that simulates a random walk. Instead of distributing chips randomly, each position makes the chips move according to the walk's possible steps in a fixed order. A random walk is called *Proppian* if at each time at each position the number of chips differs from the expected value by at most a constant, independent of time or the initial configuration of chips.

The simple walk where the possible steps are 1 or -1 each with probability $p = \frac{1}{2}$ is Proppian, with constant approximately 2.29. The equivalent simple walks on $\mathbb{Z}^d$ are also Proppian. Here, we show the same result for a larger class of walks on $\mathbb{Z}$, allowing an arbitrary number of possible steps with some constraint on their probabilities.

1. Introduction

Consider a random walk W_t on $\mathbb{Z}$ such that in each time step there are ℓ possible moves with different probabilities $\Pr[W_{t+1} - W_t = a_k] = p_k$ for $k = 1$ to ℓ, such that $\sum_{i=1}^{\ell} a_i p_i = 0$. Let σ^2 denote the variance of $W_{t+1} - W_t$. For convenience, we further assume without loss of generality that $|a_1| \geq |a_2| \geq \cdots \geq |a_\ell|$. Assume that all the p_i in the random walk are rational numbers. Also assume that $\gcd(\{a_k - a_1\}_{k=2}^{\ell}) = 1$. This makes the walk *aperiodic*: for every position x, there is a time t_x so that for every $t > t_x$, position x is reachable from the origin in t steps.

Consider an arbitrary number of *chips* independently following the same random walk. Let $L(x, t)$ denote the expected number of chips at position x at time t. For $t = 0$, there is some initial configuration $L(\cdot, 0)$. For $t > 0$, $L(x, t)$ is given by the recursive formula $L(x, t) = \sum_{k=1}^{\ell} L(x - a_k, t - 1)p_k$. The *Linear Machine* is a simulation of this random walk.

Definition 1.1. $E(x, t)$ is the number of chips at position x after running the Linear Machine for t time units from a given initial configuration.

We can think of $E(y, t)$ as some number of chips that are in position y. In the *Linear Machine*, critically, the number of chips at a given position need not be integral. At each time unit, this machine sends a proportion of $E(y, t)$ to each of the $y + a_k$ with their corresponding probabilities p_k. At any given time, the number of chips in a given position represents the expected number of chips at that position if each chip in the initial configuration were following the random walk independently, that is, $E \equiv L$.

For example, if $\ell = 2$ and $a_1 = -2$, $a_2 = 1$, then $p_1 = 1/3$ and $p_2 = 2/3$. If there are three chips at a given position, one will go to the left and two will go to the right. But if there is only one chip, then one third of that chip will go to the left and two thirds will go to the right. We use this example to describe the machine for simplicity, but it is not aperiodic. One aperiodic example would be $\ell = 3,\ a_1 = -2,\ a_2 = -1,\ a_3 = 1,\ p_1 = 1/9,\ p_2 = 1/3,\ p_3 = 5/9$.

* Corresponding author. Tel.: +55 21 91364933; fax: +55 21 35271282.
E-mail addresses: jufreire@gmail.com (J. Freire), spencer@cims.nyu.edu (J. Spencer).
1 Tel.: +1 212 998 3219; fax: +1 212 995 4124.

0012-365X/$ – see front matter
doi:10.1016/j.disc.2010.11.001

The *Propp Machine* is another way to simulate the random walk, but in such a way that the number of chips in position y at time t is always an integer. In other words, it is not possible for a portion of a chip to move to each of the different positions. One of the $y + a_k$ will be chosen to receive the whole chip.

Informally, the Propp Machine will work as follows: beginning from an initial configuration of some chips in given positions, each position will distribute those chips in numbers proportional to the probabilities of the different steps whenever the number of chips in that position makes that possible. For example, if $\ell = 2$ and $a_1 = -2$, $a_2 = 1$, then $p_1 = 1/3$ and $p_2 = 2/3$. If there are three chips at a given position, one will go to the left and two will go to the right. But if there are four chips, three of them will be treated as in the previous case, and there is an extra chip. That chip will be sent to one of the $y + a_k$ chosen by the Propp Machine according to a fixed, pre-defined sequence S_j. Each position will follow this sequence independently, keeping track of what the next choice will be by means of an "arrow" $A(y, t)$ which is updated every time the Propp Machine sends an extra chip somewhere. The initial arrow $A(y, 0)$ can be arbitrary. The sequence S_j contains step a_k a number of times that is proportional to its probability in the random walk. In this example it would be $-2, 1, 1, -2, 1, 1, \ldots$.

The Propp Machine cannot cut chips in half, so at any single time, the chips at a given y may move to new positions in a ratio that does not accurately reflect the probabilities p_j. Critically, the arrow at y "remembers" this uneven distribution and rebalances at later times, by sending the next extra chip to the next S_j.

In this example, start with, say, $A(0, 0) = 0$. If there are four chips at $y = 0$, $t = 0$, the extra one will be sent to the left and $A(0, 1) = 1$. Then at $t = 1$ if there is an extra chip it will be sent to the right and $A(0, 2) = 2$. At time $t = 2$, if there are two extra chips, one will go to the right and one to the left, and $A(0, 3) = 4$.

The set of elements of S_j used by each position at each time will be called $\mathrm{POS}[x, t]$. For example, we could have $\mathrm{POS}[y, 0] = \{4, 5, 6\}$, $\mathrm{POS}[y, 1] = \emptyset$, $\mathrm{POS}[y, 2] = \{7, 8, 9, 10\}$, $\mathrm{POS}[y, 3] = \{11\}$, $\mathrm{POS}[y, 4] = \emptyset$, $\mathrm{POS}[y, 5] = \{12, 13, 14, 15\}$.

More formally, we write $f(x, t)$ for the number of chips at location x at time t on the Propp Machine, and define the initial configuration and initial arrow as follows.

Definition 1.2. $f(\cdot, 0)$ is any function from $\mathbb{Z}$ to $\mathbb{N}$ such that $\sum_j f(j, 0)$ is finite.

Definition 1.3. $A(\cdot, 0)$ is any function from $\mathbb{Z}$ to $\mathbb{N}$.

Let $p_i = q_i/r_i$ and $r = \mathrm{lcm}(\{r_i\}_{i=1}^{\ell})$.

Definition 1.4. S_j, $0 \le j < r$, is any sequence such that for $n_i = p_i r$ values of j, $S_j = a_i$. Extend the sequence S_j periodically to all $j \ge 0$, setting $S_{rq+j} = S_j$ for all $0 \le j < r$, $q \ge 1$.

The restriction that the p_i are rational guarantees that we can make such a finite sequence S_j.

For each t, we now define the Propp Machine recursively.

Definition 1.5. $A(x, t) = \max(\{A(x, 0) - 1\} \cup \bigcup_{0 \le s < t} \mathrm{POS}[x, s]) + 1$.

Definition 1.6. $\mathrm{POS}[x, t] = \{j : A(x, t) \le j < A(x, t) + f(x, t) - \lfloor f(x, t)/r \rfloor r\}$.

Definition 1.7.

$$f(x, t + 1) = \sum_{j=1}^{\ell} [\lfloor f(x - a_j, t) p_j \rfloor + |\{i \in \mathrm{POS}[x - a_j, t] : S_i = a_j\}|].$$

The Propp Machine is also known as the rotor-router model, and has been explored extensively. See, for example, the survey in [5]. For another recent description of the rotor-router model, see [6].

From the description of the Linear Machine and the Propp Machine above, it is clear that the integer $f(x, t)$ cannot always equal $E(x, t)$.

Definition 1.8. A random walk is *Proppian* if the difference between $f(x, t)$ and $E(x, t)$ is bounded by a constant not depending on the number of chips, the initial configuration, the initial arrows, x, or t.

Remark 1.9. For an initial configuration with just one chip, $|E(x, t) - f(x, t)| \le 1$. For any initial configuration, at time $t = 1$, $|E(x, 1) - f(x, 1)| \le \ell$ as errors at x have come from the ℓ roundoffs. The strength of the definition of Proppian is that the constant is independent of both the number of chips and the time t.

Our result for $\mathbb{Z}$ can be stated as follows.

Theorem 1.10. *For a random walk with steps $a_1, \ldots, a_\ell$ and respective probabilities $p_1, \ldots, p_\ell$ satisfying $\sum_{i=1}^{\ell} a_i p_i = 0$ (no drift) and $\gcd(\{a_k - a_1\}_{k=2}^{\ell}) = 1$ (aperiodic), there is a constant C depending on the steps $a_1, \ldots, a_\ell$ and the probabilities $p_1, \ldots, p_\ell$, but not depending on t nor on the initial configuration or arrows, such that*

$$|f(0, t) - E(0, t)| < C.$$

Cooper et al. [2] showed among other results that this is true for the case $l = 2, a_1 = 1, a_2 = -1$, with constant $C = 2.29$, when the starting configuration has chips only on even positions. Here, we allow any starting configuration by restricting our attention to aperiodic walks, but a similar result would hold for periodic walks as long as we were careful with the initial configuration.

The notion of the Propp Machine and of Proppian naturally extends to any graph for which all vertices have finite degree and any random walk with rational probabilities on that graph. Cooper and Spencer [3] also showed that the simple random walk in $\mathbb{Z}^d$ is Proppian. However, there are random walks on other graphs which are not Proppian, as shown by Cooper et al. [1]. A full characterization of Proppian random walks remains elusive.

2. Outline of the proof

Definition 2.1. $H(x, t)$ is the probability that a chip at position x is at the origin in t steps when following a random walk with the given probabilities.

Definition 2.2. $E(x, t_1, t_2)$ is the number of chips at position x after running the Propp Machine for t_1 steps, and then running the Linear Machine for $t_2 - t_1$ steps.

Definition 2.3. $\mathrm{INFL}(j; x, t) = H(x + S_j, t - 1) - H(x, t)$ is the difference in the expected final number of chips at the origin when we move a single chip at time t from position x to $x + S_j$ in a step of the Propp Machine instead of sending portions of that chip to the various $x + a_k$ by a step of the Linear Machine.

The difference in Theorem 1.10 can be split into a sum of terms

$$f(0, t) - E(0, t) = \sum_{s=0}^{t-1} (E(0, s + 1, t) - E(0, s, t)).$$

For each time s, the difference $E(0, s + 1, t) - E(0, s, t)$ between the Propp step and the linear step can be further split into a sum of this difference for each chip that acts as an "extra" chip at that time:

$$E(0, s + 1, t) - E(0, s, t) = \sum_{y \in \mathbb{Z}} \sum_{j \in \mathrm{POS}[y,s]} \mathrm{INFL}(j; y, t - s),$$

which adds the correct influence for each extra chip, since for each y, one $\mathrm{INFL}(j; \cdot, \cdot)$ is added for each $j \in \mathrm{POS}[y, s]$, which are the steps that were chosen for those extra chips. Now we have

$$f(0, t) - E(0, t) = \sum_{s=0}^{t-1} \sum_{y \in \mathbb{Z}} \sum_{j \in \mathrm{POS}[y,s]} \mathrm{INFL}(j; y, t - s).$$

Changing the order of the sums, we get

$$f(0, t) - E(0, t) = \sum_{y \in \mathbb{Z}} \sum_{s=0}^{t-1} \sum_{j \in \mathrm{POS}[y,s]} \mathrm{INFL}(j; y, t - s).$$

To prove Theorem 1.10, we show

Theorem 2.4. *For a fixed y,*

$$\left| \sum_{s=0}^{t-1} \sum_{j \in \mathrm{POS}[y,s]} \mathrm{INFL}(j; y, t - s) \right| = O(y^{-2}).$$

The constant implicit in $O(y^{-2})$ depends only on the a_i and p_i.

We shall show Theorem 2.4 by first proving the equivalent statement for an approximation of $\mathrm{INFL}(j; y, t)$ (Theorem 3.1). We shall then show that the errors in the approximation do not accumulate (Theorems 4.2 and 5.7).

Definition 2.5.

$$G(y, t) = \frac{e^{-y^2/2\sigma^2 t}}{\sqrt{2\pi t}\sigma}.$$

$G(y, t)$ provides a natural approximation for $H(x, t)$, as a walk of t steps is nearly Gaussian. We write

$$H(y, t) = G(y, t) + \mathrm{ERROR}_G(y, t).$$

Since

$$H(y,t)=\sum_{j=1}^{\ell} p_j H(y+a_j,t-1),$$

we have

$$\mathrm{INFL}(j;y,t)=H(y+a_j,t-1)-\sum_{k=1}^{\ell} p_k H(y+a_k,t-1). \tag{1}$$

It is valuable to write $\mathrm{INFL}(j;y,t)$ in this way as all the terms are now referring to the same time $t-1$.

Definition 2.6.

$$\mathrm{INFL}_G(j;y,t)=G(y+a_j,t-1)-\sum_{k=1}^{\ell} p_k G(y+a_k,t-1).$$

Now

$$\begin{aligned} f(0,t)-E(0,t) &= \sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in \mathrm{POS}[y,s]} \mathrm{INFL}_G(j;y,t-s) \\ &\quad +\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in \mathrm{POS}[y,s]} \mathrm{ERROR}_G(y+a_j,t-s-1) \\ &\quad -\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in \mathrm{POS}[y,s]}\sum_{k=1}^{\ell} p_k\,\mathrm{ERROR}_G(y+a_k,t-s-1). \end{aligned} \tag{2}$$

We can also make use of an approximation by Taylor series,

$$G(y+a_j,t)=G(y,t)+a_j\frac{\mathrm{d}}{\mathrm{d}y}G(y,t)+\mathrm{ERROR}_T(j;y,t), \tag{3}$$

where

$$\mathrm{ERROR}_T(j;y,t)=\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}y^2}G(z,t)$$

for some $z\in(y,y+a_j)$. This means that

$$\mathrm{INFL}_G(j;y,t)=a_j\frac{\mathrm{d}}{\mathrm{d}y}G(y,t-1)+\mathrm{ERROR}_T(j;y,t-1)-\sum_{k=1}^{\ell} p_k\,\mathrm{ERROR}_T(k;y,t-1). \tag{4}$$

Combining (2) and (4) we get

$$f(0,t)-E(0,t) = \sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in \mathrm{POS}[y,s]} a_j\frac{\mathrm{d}}{\mathrm{d}y}G(y,t_i) \tag{5}$$

$$+\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in \mathrm{POS}[y,s]} \mathrm{ERROR}_T(j;y,s) \tag{6}$$

$$-\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in \mathrm{POS}[y,s]}\sum_{k=1}^{\ell} p_k\,\mathrm{ERROR}_T(k;y,s) \tag{7}$$

$$-\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in \mathrm{POS}[y,s]}\sum_{k=1}^{\ell} p_k\,\mathrm{ERROR}_G(y+a_k,s) \tag{8}$$

$$+\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in \mathrm{POS}[y,s]} \mathrm{ERROR}_G(y+a_j,s). \tag{9}$$

The approach used in [2] is to show that for a fixed y, the influence function is unimodal as a function of t. But with $\ell > 2$, in general $H(x,t)$ cannot be written in closed form. Instead of analyzing the unimodality of the influence function, in Section 3 we will use the fact that $\frac{\mathrm{d}}{\mathrm{d}y}G(y,t)$ is unimodal as a function of t to bound the approximation (5) by $|\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in\mathrm{POS}[y,s]} a_j \frac{\mathrm{d}}{\mathrm{d}y}G(y,t_i)| < C_a$, where C_a depends only on a_i and p_i.

In Section 4 we show that $|\sum_{y\in\mathbb{Z}}\sum_s \mathrm{ERROR}_T(y,s)| < C_t$, where ERROR_T can be any of the $\mathrm{ERROR}_T(j;\cdot,\cdot)$ and C_t depends only on a_i and p_i. On line (6), we can split the sum into ℓ parts, each one with fixed j:

$$\begin{aligned}\left|\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in\mathrm{POS}[y,s]}\mathrm{ERROR}_T(j;y,s)\right| &\le \sum_{j=1}^{\ell}\left|\sum_{y\in\mathbb{Z}}\sum_{s;j\in\mathrm{POS}[y,s]}\mathrm{ERROR}_T(j;y,s)\right|\\ &\le \sum_{j=1}^{\ell}\left|\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\mathrm{ERROR}_T(j;y,s)\right|\\ &\le \ell C_t.\end{aligned}$$

On line (7),

$$\left|\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in\mathrm{POS}[y,s]}\sum_{k=1}^{\ell}p_k\,\mathrm{ERROR}_T(k;y,s)\right| \le \sum_{k=1}^{\ell}p_k\left|\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in\mathrm{POS}[y,s]}\mathrm{ERROR}_T(k;y,s)\right| \le \ell C_t.$$

In Section 5 we show $|\sum_{y\in\mathbb{Z}}\sum_s \mathrm{ERROR}_G(y,s)| < C_g$, where C_g depends only on a_i and p_i. Similarly to what we do for (7) and (6), lines (8) and (9) can be rearranged. By using $\tilde{y} = y + a_k$ or $\tilde{y} = y + a_j$, we can use the same upper bound of C_g for each of the new sums.

The proof is then completed by setting $C = C_a + 2\ell C_t + 2\ell C_g$.

3. Approximation by Gaussian

In this section we are going to show the following.

Theorem 3.1. *Consider an aperiodic random walk with possible steps $a_1, a_2, \ldots, a_\ell$ with probabilities $p_1, p_2, \ldots, p_\ell$, such that $\sum a_i p_i = 0$. Then*

$$\left|\sum_{y\in\mathbb{Z}}\sum_{i\ge A(y,0)} S_i\frac{\mathrm{d}}{\mathrm{d}y}G(y,t_i)\right| < C_a = \frac{r\sum_{j=1}^{\ell}|p_j a_j|}{2}\,\frac{\sqrt{3}\mathrm{e}^{-3/2}\pi^2}{\sqrt{2\pi}},$$

for a sequence of positive terms $t_{i+1} \ge t_i$.

Because we allow $t_{i+1} = t_i$, this theorem will imply that

$$\left|\sum_{y\in\mathbb{Z}}\sum_{s=0}^{t-1}\sum_{j\in\mathrm{POS}[y,s]} a_j\frac{\mathrm{d}}{\mathrm{d}y}G(y,t_i)\right| < C_a.$$

We will have $t_i = s$ for a number of i that is the same as the size of the set $\mathrm{POS}[y,s]$, and one S_j appears for each of the $j \in \mathrm{POS}[y,s]$.

Proof. We start by looking at

$$\frac{\mathrm{d}}{\mathrm{d}y}G(y,t) = \frac{\mathrm{e}^{-y^2/2t\sigma^2}}{\sigma\sqrt{2\pi t}}\left(-\frac{y}{t\sigma^2}\right)$$

for $t > 0$. For a fixed y, this is unimodal as a function of t, as it has only one maximum at

$$\frac{\mathrm{d}}{\mathrm{d}y}G\left(y,\frac{y^2}{3\sigma^2}\right) = \frac{3^{3/2}\mathrm{e}^{-3/2}}{\sqrt{2\pi}y^2}. \tag{10}$$

Now we look again at

$$\sum_{i\ge A(y,0)} S_i\frac{\mathrm{d}}{\mathrm{d}y}G(y,t_i)$$

for a fixed y. It is important that the S_i were chosen to have the right proportions of each a_k.

Lemma 3.2. *Let α_j be a sequence of integers with period r such that $\sum_{j=0}^{r-1} \alpha_j = 0$. Consider a sequence $0 < t_0 \le t_1 \le \cdots$ such that $t_j \to \infty$ as $j \to \infty$. Let $F(t)$ be a unimodal function such that $F(t) \to 0$ as $t \to \infty$ and such that the sum $\sum_j \alpha_j F(t_j)$ is absolutely convergent. Then*

$$\left| \sum_j \alpha_j F(t_j) \right| < C \max_t |F(t)|.$$

Proof. The first r terms of the sum are

$$\sum_{j=0}^{r-1} \alpha_j F(t_j) = \sum_{j=0}^{r-1} \sum_{k=1}^{|\alpha_j|} \pm F(t_j),$$

which is a sum of $\sum_{j=0}^{r-1} |\alpha_j|$ terms $\pm F(t_j)$, with exactly $\sum_{j=0}^{r-1} |\alpha_j|/2$ being positive and $\sum_{j=0}^{r-1} |\alpha_j|/2$ being negative.

Since $F(t_j)$ is unimodal, a sum with alternating signs is bounded by the maximum of its absolute value. The sequence of signs in the sum can be broken down into at most $\sum_{j=0}^{r-1} |\alpha_j|/2$ alternating sequences. Therefore,

$$\left| \sum_j \alpha_j F(t_j) \right| < \frac{\max_t |F(t)|}{2} \sum_{j=0}^{r-1} |\alpha_j|. \quad \square$$

We know that

$$\sum_{j=0}^{r-1} S_j = \sum_{k=1}^{\ell} n_k a_k = 0$$

because $n_k = p_k r$, so we can set $\alpha_j = S_j$. Since $\frac{\mathrm{d}}{\mathrm{d}y} G(y,t)$ is unimodal in t and goes to zero as $t \to \infty$, we can set $F(t) = \frac{\mathrm{d}}{\mathrm{d}y} G(y,t)$ for a fixed y. The sequence t_j will be the times t for which POS$[y,t]$ is not empty, repeated as many times as there are elements in POS$[y,t]$. Therefore, t_j is either a finite sequence or a sequence with $t_j \to \infty$, and the sum in Lemma 3.2 is absolutely convergent.

For example, take $\ell = 2$, $a_1 = -1$, $a_2 = 2$, $p_1 = 2/3$ and $p_2 = 1/3$. The arrow sequence S_j will be $S_0 = -1$, $S_1 = -1$, $S_2 = 2$. The sequence of signs will be $--++--++\cdots$ which is two alternating sequences.

The choice of S_j is relevant to the final constant in the bound, as shown (for $\mathbb{Z}^2$) in [4]. In our work, however, we do not attempt to find the best constants.

Returning to Theorem 3.1, we now have the upper bound

$$\begin{aligned} \left| \sum_{y\in\mathbb{Z}} \sum_{i\ge A(y,0)} S_i \frac{\mathrm{d}}{\mathrm{d}y} G(y,t_i) \right| &\le \frac{\sum_{j=0}^{r-1} |S_j|}{2} \frac{3^{3/2}\mathrm{e}^{-3/2}}{\sqrt{2\pi}} \sum_{y\in\mathbb{Z}} \frac{1}{y^2} \\ &= \frac{r\sum_{j=1}^{\ell} |p_j a_j|}{2} \frac{3^{3/2}\mathrm{e}^{-3/2}}{\sqrt{2\pi}} \frac{\pi^2}{3} \\ &= C_a. \end{aligned}$$

Notice that we ignored the term $y = 0$ in the sum, which would be an infinite term. The case $y = 0$ can be treated separately. The quantity INFL(y,t) is only defined for $t \ge 1$. For $t \ne 0$, $G(0,t) = \frac{1}{\sigma\sqrt{2\pi t}}$ and $\frac{\mathrm{d}}{\mathrm{d}y} G(0,t) = 0$. So the contribution of position $y = 0$ is zero. $\square$

4. Error from Taylor series

The function $H(y,t)$ represents the probability that the walk gets to the origin from y at time t. We have approximated $H(y,t)$ by $G(y,t)$. There is a question of how good this approximation is, which we will deal with in the next section. For now, consider the approximation.

Definition 4.1. Without loss of generality, assume $a > 0$.

$$\mathrm{ERROR}_T(y,t) = \frac{\mathrm{d}^2}{\mathrm{d}y^2} G(z,t) = 2\frac{\mathrm{d}}{\mathrm{d}t} G(z,t),$$

for some $z \in [y, y+a]$.

This error appears in the approximation (3), where a can be any of the a_j.

Theorem 4.2.

$$\left|\sum_{y\in\mathbb{Z}}\sum_{i\geq 0}\mathrm{ERROR}_T(y,t_i)\right| < C_t,$$

where the constant C_t depends only on a_i and p_i.

Proof. For convenience, we will use $\frac{\mathrm{d}}{\mathrm{d}t}G(y,t)$ instead of $\frac{\mathrm{d}}{\mathrm{d}t}G(z,t)$, where z may vary for each term. This will be justified shortly. The error term is

$$\frac{\mathrm{d}}{\mathrm{d}t}G(y,t) = \frac{\mathrm{e}^{-y^2/2\sigma^2 t}}{2\sqrt{2\pi}\sigma^3 t^{5/2}}(y^2-\sigma^2 t).$$

For a fixed y, this is zero for $t = t_0 = y^2/\sigma^2$ and for $t = 0$, and it goes to zero when $t \to \infty$. There are two maxima in absolute value, at

$$t_1 = \frac{y^2(6-\sqrt{24})}{6\sigma^2}$$

and

$$t_2 = \frac{y^2(6+\sqrt{24})}{6\sigma^2}.$$

We have

$$\frac{\mathrm{d}}{\mathrm{d}t}G(y,t_1) = O(y^{-3}) \tag{11}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}G(y,t_2) = O(y^{-3}). \tag{12}$$

One possible approach would be to bound the sum by an integral. For a given y,

$$\begin{aligned}\sum_{i\geq 0}\frac{\mathrm{d}}{\mathrm{d}t}G(y,t_i) &\leq \int_0^{t_0}\frac{\mathrm{d}}{\mathrm{d}t}G(y,t)\mathrm{d}t + \int_{t_0}^{\infty}-\frac{\mathrm{d}}{\mathrm{d}t}G(y,t)\mathrm{d}t + 2\left|\frac{\mathrm{d}}{\mathrm{d}t}G(y,t_1)\right| + 2\left|\frac{\mathrm{d}}{\mathrm{d}t}G(y,t_2)\right| \\ &= 2\left|\frac{\mathrm{d}}{\mathrm{d}t}G(y,t_1)\right| + 2\left|\frac{\mathrm{d}}{\mathrm{d}t}G(y,t_2)\right| + 2G(y,t_0) - G(y,0) - \lim_{t\to\infty}G(y,t) \\ &= \frac{c_1}{y^3} + \frac{c_2}{y^3} + \frac{c_3}{y} + 0 + 0,\end{aligned}$$

but this is not small enough, because of the term $\frac{c_3}{y}$. This term came from the integration close to t_0. In this range, however, we will see that INFL_G is itself unimodal, so that the approximation by $\frac{\mathrm{d}}{\mathrm{d}y}G(y,t)$ and the sum of the errors are not necessary. Putting (10)–(12) together with (4), we find that

$$\max_t \mathrm{INFL}_G(y,t) = \Theta(y^{-2}).$$

This changes the value of the constant C_a, but it still depends only on the a_j and p_j.

To find out in what ranges of t $\mathrm{INFL}_G(y,t)$ is unimodal, we must look at $\Delta\mathrm{INFL}_G(y,t) = \mathrm{INFL}_G(y,t+1) - \mathrm{INFL}_G(y,t)$.

$$\mathrm{INFL}_G(y,t+1) = \mathrm{INFL}_G(y,t) + \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{INFL}_G(y,t) + \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{INFL}_G(y,s),$$

for some $s \in [t, t+1]$.

$$\Delta\mathrm{INFL}_G(y,t) = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{INFL}_G(y,t) + \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{INFL}_G(y,s),$$

and since

$$\mathrm{INFL}_G(y,t) = \frac{\mathrm{d}}{\mathrm{d}y}G(y,t) + \frac{\mathrm{d}^2}{\mathrm{d}y^2}G(z,t),$$

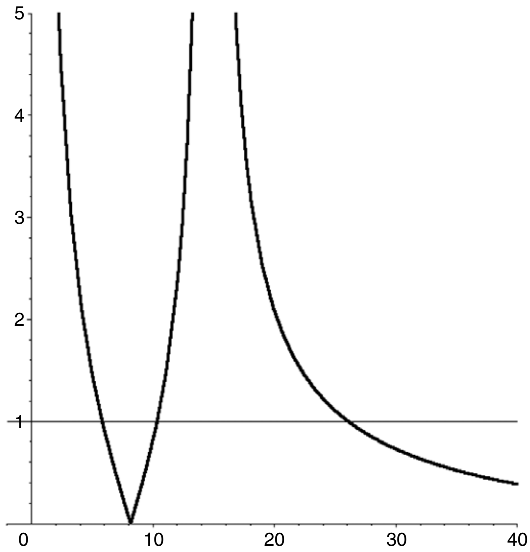

Fig. 1. Ratio of $\frac{d^3}{dy^3}G(y,t)$ to $\frac{d^4}{dy^4}G(y,t)$. $y = 20$, $\sigma = 3$.

we have

$$\triangle\mathrm{INFL}_G(y,t) = \frac{1}{2}\frac{d^3}{dy^3}G(y,t) + \frac{1}{2}\frac{d^4}{dy^4}G(z,t) + \frac{1}{8}\frac{d^5}{dy^5}G(y,s) + \frac{1}{8}\frac{d^6}{dy^6}G(z,s). \tag{13}$$

The first term, $\frac{1}{2}\frac{d^3}{dy^3}G(y,t)$, is the approximation $\triangle\mathrm{INFL}_G(y,t) \sim \frac{d}{dt}\frac{d}{dy}G(y,t)$. If the error in the approximation of $\triangle\mathrm{INFL}_G(y,t)$ is smaller in absolute value than the approximation itself, then we can say that the approximation has the right sign. If that is the case, then $\mathrm{INFL}_g(y,t)$ and $\frac{d}{dy}G(y,t)$ are both increasing or decreasing together, so if one is unimodal, the other one is also unimodal. The sum of the three error terms in (13) needs to be less than the approximation. It is sufficient that the absolute value of each error term be less than one third of the approximation. For the first of the error terms, we need the ratio

$$\frac{1}{t\sigma^2 y}\frac{6(y^4 - 6y^2 t\sigma^2 + 3t^2\sigma^4)}{y^2 - 3t\sigma^2} < 1.$$

Fig. 1 is a picture of the absolute value of the ratio, for $y = 20$ and $\sigma = 3$. It goes to infinity when $t = 0$ and $t = y^2/3\sigma^2$, which is the point of maximum of $\mathrm{INFL}_G(y,t)$.

In the first of these ranges, choosing $t = \omega(y)$ makes the ratio go to zero as $y \to \infty$. Take, for specificity, $t = y\log(y)$, though a wide range of $t = t(y)$ would work. There is a y_0 such that the ratio is < 1 for all $y > y_0$ and $t > y\log(y)$. This range also solves the problem for the other two error terms.

The sum of ERROR_T from $t = 0$ to $y\log(y)$ is still needed. This is bounded by

$$\int_0^{y\log(y)} \frac{d}{dt}G(z,t)dt$$

where $z \in (y, y+a)$ maximizes $\frac{d}{dt}G(z,t)$ in that interval for a given t. In principle, it could be a different z for each t in the sum, but $\frac{d}{dt}G(y,t)$ is increasing with y only for $y < \sqrt{3\sigma^2 t}$. For $t \in [0, y\log(y)]$, $\frac{d}{dt}G(y,t)$ is decreasing with y, so the maximizing z is actually $z = y$. Now

$$\int_0^{y\log(y)} \frac{d}{dt}G(y,t)dt = G(y, y\log(y)) = \frac{e^{-y/2\log(y)\sigma^2}}{\sqrt{2\pi}\sigma\sqrt{y\log(y)}} = o(y^{-2}).$$

The other bad range is at $t = y^2/3\sigma^2$. The value of the first ratio at $t = y^2/3\sigma^2 \pm ay/\sigma^2$ is $\Theta(1)$, and less than 1 if $a > 4$. There is a y_0 such that this ratio is < 1 for all $y > y_0$. The same thing is true for the other two ratios.

Summing the errors in this range, we get, setting $a = 5$,

$$\int_{y^2/3\sigma^2 - 5y/\sigma^2}^{y^2/3\sigma^2 + 5y/\sigma^2} \frac{d}{dt}G(z,t)dt \le 10\frac{y}{\sigma^2}\max_t\left(\max_{z\in(y,y+a)}\frac{d}{dt}G(z,t)\right) = \Theta(y^{-2}). \quad \square$$

5. Error from approximation by Gaussian

We begin this section with a lemma stating that no approximations, not even by a Gaussian, are necessary for $t \le y$.

Lemma 5.1. *For* INFL(y, t) *the influence of one step of type* a_1,

$$\sum_{t \le y} |\text{INFL}(y,t)| \le \beta y e^{\alpha y},$$

for constants $\alpha < 0$ *and* β *that depend only on the* a_i *and* p_i*. In particular, they do not depend on* y.

Proof. For $t \le y$,

$$|\text{INFL}(y,t)| = |H(y-a_1, t-1) - H(y,t)| \le |H(y-a_1,t-1)| + |H(y,t)|,$$

and

$$\begin{aligned} H(y,t) &= \Pr[X_1 + \cdots + X_t = y] \\ &\le E[e^{\lambda X_1}]^t e^{-\lambda y}, \end{aligned}$$

by applying Markov's inequality to the quantity $e^{\lambda \sum_{i=1}^t X_i}$ and using the fact that the X_i are independent.

Since X_1 takes on a finite number of values, there is an M such that $|X_1| < M$. We can therefore bound $|X_i^m| < M^{m-2} X_i^2$ for $m \ge 3$. We bound the Laplace Transform (noting that $E[X] = 0$)

$$\begin{aligned} E[e^{\lambda X_1}] &= E\left[1 + \lambda X_1 + \frac{\lambda^2 X_1^2}{2} + \frac{\lambda^3 X_1^3}{3!} + \cdots\right] \\ &\le E\left[1 + \lambda X_1 + \frac{\lambda^2 X_1^2}{2}\left(1 + \frac{\lambda M}{1!} + \frac{\lambda^2 M^2}{2!} + \cdots\right)\right] \\ &= E\left[1 + \lambda X_1 + \frac{\lambda^2 X_1^2}{2} e^{\lambda M}\right] \\ &= 1 + \frac{\lambda^2 \sigma^2}{2} e^{\lambda M} \\ &\le e^{\frac{\lambda^2 \sigma^2}{2} e^{\lambda M}}. \end{aligned}$$

Now we can say that

$$\begin{aligned} H(y,t) &\le E[e^{\lambda X_1}]^t e^{-\lambda y} \\ &\le e^{\frac{\lambda^2 \sigma^2}{2} e^{\lambda M} t - \lambda y}. \end{aligned}$$

Choose a constant λ such that $\lambda e^{\lambda M} \le \frac{2}{\sigma^2}$. This is possible because $\lambda e^{\lambda M} \xrightarrow[\lambda \to 0]{} 0$. Then $\frac{\lambda^2 \sigma^2}{2} e^{\lambda M} < \lambda$ and since $t \le y$,

$$H(y,t) \le e^{\alpha y},$$

where $\alpha = \frac{\lambda^2 \sigma^2}{2} e^{\lambda M} - \lambda < 0$, so that $H(y,t)$ is exponentially decreasing in y.

Also,

$$\begin{aligned} H(y-a_1, t-1) &\le e^{\frac{\lambda^2 \sigma^2}{2} e^{\lambda M}(t-1) - \lambda(y-a_1)} \\ &= e^{\alpha y} e^{\lambda a_1}, \end{aligned}$$

and

$$\begin{aligned} |\text{INFL}(y,t)| &\le |H(y-a_1,t-1)| + |H(y,t)| \\ &\le (1 + e^{\lambda a_1}) e^{\alpha y} \\ &= \beta e^{\alpha y}. \end{aligned}$$

Finally,

$$\sum_{t \le y} |\text{INFL}(y,t)| \le \beta y e^{\alpha y}. \quad \square$$

Lemma 5.2.

$$\sum_{y \in \mathbb{Z}} \sum_{i: t_i < y} \text{INFL}(y, t_i) < \infty.$$

Proof. This follows directly from the previous lemma.

$$\left|\sum_{i:t_i\le y} \mathrm{INFL}(y,t_i)\right| \le \sum_{t\le y} |\mathrm{INFL}(y,t)| \le \beta y e^{\alpha y} = o(y^{-2}),$$

and

$$\sum_y \beta y e^{\alpha y} < \infty. \quad \square$$

Definition 5.3. ERROR_G is the error in the approximation

$$H(y,t) = G(y,t) + \mathrm{ERROR}_G(y,t). \tag{14}$$

We know (see [7], Theorem 2.3.5) that

$$\left| H(y,t) - \frac{1}{\sigma\sqrt{2\pi t}} e^{-y^2/2\sigma^2 t}\right| < \frac{c_1}{t},$$

where c_1 depends only on the a_i and p_i but is uniform for all t, y. This error is too big, but from Theorem 2.3.8 in [7] there is also a sequence of better approximations with smaller errors.

Definition 5.4.

$$\mathrm{APPROX}(y,t) = e^{-y^2/2\sigma^2 t}\left(\frac{1}{\sigma\sqrt{2\pi t}} + \frac{u_3(z)}{t} + \frac{u_4(z)}{t^{3/2}} + \frac{u_5(z)}{t^2} + \frac{u_6(z)}{t^{5/2}} + \frac{u_7(z)}{t^3}\right), \tag{15}$$

where $z = y/\sqrt{t}$ and $u_n(z)$ is a polynomial of degree n.

Writing out all the terms in the polynomials u_n, this approximation is

$$\begin{aligned} e^{-y^2/2\sigma^2 t}\Bigg(&\frac{1}{\sigma\sqrt{2\pi t}} + \frac{c_{3,0}}{t} + \frac{c_{3,1}y}{t^{3/2}} + \frac{c_{3,2}y^2}{t^2} + \frac{c_{3,3}y^3}{t^{5/2}} + \frac{c_{4,0}}{t^{3/2}} + \frac{c_{4,1}y}{t^2} + \frac{c_{4,2}y^2}{t^{5/2}} + \frac{c_{4,3}y^3}{t^3} + \frac{c_{4,4}y^4}{t^{7/2}} \\ &+ \frac{c_{5,0}}{t^2} + \frac{c_{5,1}y}{t^{5/2}} + \frac{c_{5,2}y^2}{t^3} + \frac{c_{5,3}y^3}{t^{7/2}} + \frac{c_{5,4}y^4}{t^4} + \frac{c_{5,5}y^5}{t^{9/2}} + \frac{c_{6,0}}{t^{5/2}} + \frac{c_{6,1}y}{t^3} + \frac{c_{6,2}y^2}{t^{7/2}} + \frac{c_{6,3}y^3}{t^4} \\ &+ \frac{c_{6,4}y^4}{t^{9/2}} + \frac{c_{6,5}y^5}{t^5} + \frac{c_{6,6}y^6}{t^{11/2}} + \frac{c_{7,0}}{t^3} + \frac{c_{7,1}y}{t^{7/2}} + \frac{c_{7,2}y^2}{t^4} + \frac{c_{7,3}y^3}{t^{9/2}} + \frac{c_{7,4}y^4}{t^5} + \frac{c_{7,5}y^5}{t^{11/2}} + \frac{c_{7,6}y^6}{t^6} + \frac{c_{7,7}y^7}{t^{13/2}}\Bigg). \end{aligned} \tag{16}$$

Definition 5.5. $g_{k,m}(y,t)$ is any of the terms in (16). Notice that each of them has the form

$$g_{k,m}(y,t) = \frac{e^{-y^2/2\sigma^2 t} z^k}{t^m} = \frac{e^{-y^2/2\sigma^2 t} y^k}{t^{m+k/2}},$$

with $3 \ge m \ge 1$ and $7 \ge 2m+1 \ge k \ge 0$, where k and $2m$ are integers.

Definition 5.6. $\mathrm{DIFF}_{k,m}(y,t) = g_{k,m}(y-a_1, t-1) - g_{k,m}(y,t)$.

The main result of this section is:

Theorem 5.7.

$$\left|\sum_{y\in\mathbb{Z}}\sum_{i\ge 0} \mathrm{ERROR}_G(y,t_i)\right| < C_g,$$

where the constant C_g depends only on the a_i and p_i.

Proof. The error in the approximation $\mathrm{APPROX}(y,t)$ is

$$|H(y,t) - \mathrm{APPROX}(y,t)| < \frac{c_4}{t^{7/2}}. \tag{17}$$

Now

$$\left|\sum_{i:t_i>y} \frac{1}{t_i^{7/2}}\right| \le \left|\sum_{t=y} \frac{1}{t^{7/2}}\right| = \Theta(y^{-5/2}) = o(y^{-2})$$

and the sum of the errors is

$$\left|\sum_{y}\sum_{i:t_i>y}\frac{1}{t_i^{7/2}}\right| = \gamma < \infty.$$

However, from (14), (15) and (17), we see that

$$\mathrm{ERROR}_G(y,t) = \sum_{k,m} g_{k,m}(y,t) + \mathrm{ERROR}_A(y,t),$$

where $|\mathrm{ERROR}_A(y,t)| < c_4/t^{7/2}$ as above. Now we have to show that the terms $g_{k,m}(y,t)$ make small contributions to the sum. Namely, we want to show that

$$\sum_i (-1)^i\,\mathrm{DIFF}_{k,m}(y,t_i) = o(y^{-2}). \tag{18}$$

We have

$$\begin{aligned}\mathrm{DIFF}_{k,m}(y,t) &= g(y-a_1,t-1) - g(y,t)\\ &= -a_1\frac{\mathrm{d}}{\mathrm{d}y}g(y,t) - \frac{\mathrm{d}}{\mathrm{d}t}g(y,t) + \mathrm{ERROR}_{k,m}(y,t).\end{aligned}$$

We can write

$$\frac{\mathrm{d}}{\mathrm{d}t}g_{k,m}(y,t) = \mathrm{e}^{-y^2/2\sigma^2 t}y^k\left(\frac{y^2}{2\sigma^2 t^{2+m+k/2}} - \frac{m+k/2}{t^{m+k/2+1}}\right) \tag{19}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}y}g_{k,m}(y,t) = \frac{\mathrm{e}^{-y^2/2\sigma^2 t}}{t^{m+k/2}}\left(ky^{k-1} - \frac{y^{k+1}}{\sigma^2 t}\right), \tag{20}$$

which means that $\mathrm{DIFF}_{k,m}(y,t)$, like $\mathrm{APPROX}(y,t)$, is also a sum of terms of the form

$$T_{A,B}(y,t) = c_{A,B}\frac{\mathrm{e}^{-y^2/2\sigma^2 t}y^A}{t^B}$$

plus an error $\mathrm{ERROR}_{k,m}(y,t)$.

To show (18), we need to show

$$\sum_i (-1)^i T_{A,B}(y,t_i) = o(y^{-2}) \tag{21}$$

and

$$\sum_y\sum_i |\mathrm{ERROR}_{k,m}(y,t_i)| < \infty. \tag{22}$$

To show (21), notice that for each of the terms $T_{A,B}(y,t)$,

$$\frac{\mathrm{d}}{\mathrm{d}t}T_{A,B}(y,t) = \mathrm{e}^{-y^2/2\sigma^2 t}y^A\left(\frac{y^2}{2\sigma^2 t^{2+B}} - \frac{B}{t^{B+1}}\right),$$

which is zero only when $t = y^2/2B\sigma^2$, making $T_{A,B}(y,t)$ a unimodal function of t with maximum $\Theta(y^A/y^{2B})$.

In (19), the first term has $A = k+2$ and $B = 2+m+k/2$. The second term has $A = k$ and $B = m+k/2+1$. In both cases, $2B - A \geq 4$. In (20), there is a term with $A = k+1$ and $B = m+k/2+1$, and another term with $A = k-1$ and $B = m+k/2$. In both cases, $2B - A \geq 3$. So, the maxima of all terms $T_{A,B}(y,t)$ are $O(y^{-3})$.

The last step is to show (22). These error terms are the second derivatives of $g_{k,m}(y,t)$ taken at some point $(\tilde{y},\tilde{t})$. These are also a sum of terms $T_{A,B}(y,t)$:

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}t^2}g_{k,m}(y,t) &= \frac{\mathrm{e}^{-y^2/2\sigma^2 t}}{t^{m+k/2}}\left(\frac{y^{4+k}}{4\sigma^4 t^4} - \frac{(2+2m+k)y^{2+k}}{2\sigma^2 t^3} + \frac{((m+k/2)^2+m+k/2)y^k}{t^2}\right)\\ \frac{\mathrm{d}^2}{\mathrm{d}t\,\mathrm{d}y}g_{k,m}(y,t) &= \frac{\mathrm{e}^{-y^2/2\sigma^2 t}}{t^{m+k/2}}\left(-\frac{y^{k+3}}{2\sigma^4 t^3} + \frac{y^{k+1}(2+2k+2m)}{2\sigma^2 t^2} - \frac{y^{k-1}k(m+k/2)}{t}\right)\\ \frac{\mathrm{d}^2}{\mathrm{d}y^2}g_{k,m}(y,t) &= \frac{\mathrm{e}^{-y^2/2\sigma^2 t}}{t^{m+k/2}}\left(\frac{y^{k+2}}{\sigma^4 t^2} - \frac{y^k(1+2k)}{\sigma^2 t} + y^{k-2}(k^2-k)\right).\end{aligned}$$

A	B	$\frac{1+A}{B-1}$ (m = 1)	$\frac{2+A}{B-1}$ (m = 1)
$k+4$	$4+m+\frac{k}{2}$	$(k+5)/(\frac{k}{2}+4)$	$(k+6)/(\frac{k}{2}+4)$
$k+2$	$3+m+\frac{k}{2}$	$(k+3)/(\frac{k}{2}+3)$	$(k+4)/(\frac{k}{2}+3)$
k	$2+m+\frac{k}{2}$	$(k+1)/(\frac{k}{2}+2)$	$(k+2)/(\frac{k}{2}+2)$
$k+3$	$3+m+\frac{k}{2}$	$(k+4)/(\frac{k}{2}+3)$	$(k+5)/(\frac{k}{2}+3)$
$k+1$	$2+m+\frac{k}{2}$	$(k+2)/(\frac{k}{2}+2)$	$(k+3)/(\frac{k}{2}+2)$
$k-1$	$1+m+\frac{k}{2}$	$k/(\frac{k}{2}+1)$	$(k+1)/(\frac{k}{2}+1)$
$k+2$	$2+m+\frac{k}{2}$	$(k+3)/(\frac{k}{2}+2)$	$(k+4)/(\frac{k}{2}+2)=2$
k	$1+m+\frac{k}{2}$	$(k+1)/(\frac{k}{2}+1)$	$(k+2)/(\frac{k}{2}+1)=2$
$k-2$	$m+\frac{k}{2}$	$(k-1)/\frac{k}{2}$	$k/\frac{k}{2}=2$

Fig. 2. Table of error terms, first order approximation of $\mathrm{DIFF}(y,t)$.

We summarize these terms in the table in Fig. 2. The last two columns are filled in with the values for $m = 1$, which maximizes the terms.

We can split the sum in i into two ranges ("big" and "small"). Let the big range start at $t = y^d$, for some d to be determined. We have

$$\sum_{i:t_i\geq y^d} \frac{\mathrm{e}^{-y^2/2\sigma^2 t} y^A}{t_i^B} \leq \sum_{t\geq y^d} \frac{y^A}{t^B} = \Theta\left(\frac{y^A}{(y^d)^{B-1}}\right).$$

This is $o(y^{-2})$ iff $d(B-1)-A=\beta>2$. Notice from the table that, except for the last three lines when $m=1$, it is true that $(2+A)/(B-1)<2$. So there is a $d<2$ such that $d>(2+A)/(B-1)$, and then $d(B-1)-A=\beta>2$, and

$$\sum_{t\geq y^d} \frac{\mathrm{e}^{-y^2/2\sigma^2 t} y^A}{t^B} \leq \frac{1}{y^\beta} = o(y^{-2}).$$

Bigger d would also make this sum finite, but we need $d<2$ for the small range, as we will see later.

For the last three lines when $m=1$, we can get $\beta>1$ because $(1+A)/(B-1)<2$, and in fact we can get $\beta=2-\epsilon$ for $\epsilon>0$ as small as we want. Although this is enough for this term to be summable in y, its contribution would be bigger than the $1/y^2$ from the main term.

To get to the desired $o(y^{-2})$, we can use a better approximation for $\mathrm{DIFF}_{k,1}(y,t)$ by using more terms in the Taylor series. The next approximation is

$$\mathrm{DIFF}_{k,1}(y,t) = -a_1\frac{\mathrm{d}}{\mathrm{d}y}g(y,t) - \frac{\mathrm{d}}{\mathrm{d}t}g(y,t) + a_1^2\frac{\mathrm{d}^2}{\mathrm{d}y^2}g(y,t) + a_1\frac{\mathrm{d}^2}{\mathrm{d}y\mathrm{d}t}g(y,t) + \frac{\mathrm{d}^2}{\mathrm{d}t^2}g(y,t) + \mathrm{ERROR}_{k,1}(y,t),$$

where $\mathrm{ERROR}_{k,1}(y,t)$ are now the third derivatives.

For the terms of the second derivative we have $2B-A\geq 4$, which can be seen in the table in Fig. 2, so that

$$\sum_i (-1)^i T_{A,B}(y,t) = O(y^{-4}),$$

by using the fact that these functions are unimodal in t.

Now we look at the third derivatives, which make up the new $\mathrm{ERROR}_{k,1}$ term. We summarize the terms in the table in Fig. 3.

By the same argument as above, there is a $d<2$ such that

$$\sum_{t\geq y^d} \frac{\mathrm{e}^{-y^2/2\sigma^2 t} y^A}{t^B} \leq \frac{1}{y^\beta} = o(y^{-2}).$$

We still have to sum the errors in the small range. Notice in the table in Fig. 3 that for $m=1$ and $k=0$ we would have $B=1$ on the fourth line. This line came from a term in $\frac{\mathrm{d}^3}{\mathrm{d}y^3}g_{0,1}(y,t)$, but that term has a constant k^3-3k^2+2k. This means that the term is not present for $k=0$, so actually there is no case $B=1$.

$\mathbf{A}$	$\mathbf{B}$	$\frac{\mathbf{2+A}}{\mathbf{B-1}}$
$k+1$	$\frac{k}{2}+3$	<2
$k+3$	$\frac{k}{2}+4$	<2
$k-1$	$\frac{k}{2}+2$	<2
$k-3$	$\frac{k}{2}+1$	<2
k	$\frac{k}{2}+3$	<2
$k+2$	$\frac{k}{2}+4$	<2
k	$\frac{k}{2}+3$	<2
$k-2$	$\frac{k}{2}+2$	<2

$\mathbf{A}$	$\mathbf{B}$	$\frac{\mathbf{2+A}}{\mathbf{B-1}}$
$k+1$	$\frac{k}{2}+4$	<2
$k+3$	$\frac{k}{2}+5$	<2
$k+5$	$\frac{k}{2}+6$	<2
$k-1$	$\frac{k}{2}+3$	<2
$k+2$	$\frac{k}{2}+5$	<2
$k+4$	$\frac{k}{2}+6$	<2
$k+6$	$\frac{k}{2}+7$	<2
k	$\frac{k}{2}+4$	<2

Fig. 3. Table of error terms, second order approximation of $\mathrm{DIFF}_{k,1}(y,t)$.

Now, since we know $B > 1$ and $d = 2 - \epsilon < 2$, we have $e^{-y^2/t} \le e^{-y^\epsilon}$, and

$$\begin{aligned}\sum_{i:t_i \le y^d} T_{A,B}(y,t_i) &\le \sum_{t \le y^d} T_{A,B}(y,t)\\ &\le e^{-y^\epsilon} y^A \sum_t \frac{1}{t^B}\\ &\le e^{-y^\epsilon} y^A C = o(y^{-2}). \quad \square\end{aligned}$$

References

[1] J. Cooper, B. Doerr, T. Friedrich, J. Spencer, Deterministic random walks on regular trees, in: Proceedings of the 19th annual ACM–SIAM symposium on Discrete algorithms, 2008, pp. 766–772.

[2] J. Cooper, B. Doerr, J. Spencer, G. Tardos, Deterministic random walks on the integers, European Journal of Combinatorics 28 (8) (2007) 2072–2090.

[3] J. Cooper, J. Spencer, Simulating a random walk with constant error, Combinatorics, Probability and Computing 15 (6) (2006) 815–822.

[4] B. Doerr, T. Friedrich, Deterministic random walks on the two-dimensional grid, Combinatorics, Probability and Computing 18 (1–2) (2009) 123–144.

[5] A.E. Holroyd, L. Levine, K. Mészáros, Y. Peres, J. Propp, D.B. Wilson, Chip-firing and rotor-routing on directed graphs, in: In and Out of Equilibrium 2, in: Progr. Probab., vol. 60, Birkhäuser, 2008, pp. 331–364.

[6] A.E. Holroyd, J. Propp, Rotor walks and Markov chains, Algorithmic Probability and Combinatorics (2010) 105–126.

[7] G. Lawler, V. Limic, Random Walk: A Modern Introduction, Cambridge University Press, 2010.